Name:________________________________

MA 5230 Test 1

1. I took three tests and three pretests in my first statistics course. I was curious if my performance on the tests was related. Below you can find the scores. Please compute the r-value for me to determine if the pretest scores are coorelated with the tests scores.

Pretest score	Test score
1	5
2	9
3	10

$$\bar{X} = \frac{1+2+3}{3} = 2$$

$$S_x^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$$

$$= \frac{1}{3-1} \left[(1-2)^2 + (2-2)^2 + (3-2)^2\right]$$

$$= \frac{1}{2}\left[1 + 0 + 1\right] = 1$$

$$\boxed{S_x = 1}$$

$$\bar{Y} = \frac{5+9+10}{3} = 8$$

$$S_y^2 = \frac{1}{2}\left[(5-8)^2 + (9-8)^2 + (10-8)^2\right]$$

$$= \frac{1}{2}\left[14\right] = 7$$

$$\boxed{S_y = \sqrt{7}}$$

$$r = \frac{1}{n-1} \sum \frac{(x_i - \bar{x})}{S_x} \cdot \frac{(y_i - \bar{y})}{S_y}$$

$$r = \frac{1}{2}\left[\frac{1}{S_x \cdot S_y}\right] \cdot \left[(1-2)\cdot(5-8) + (2-2)(9-8) + (3-2)(10-8)\right]$$

$$= \frac{1}{2}\frac{1}{(1)(\sqrt{7})}\left[(-1)(-3) + 0 + (1)(2)\right]$$

$$= \frac{1}{2\sqrt{7}}\left[5\right] = \frac{5}{2\sqrt{7}}$$

$$\boxed{r = .945}$$

2. A group of 7th grade teachers want to use IQ scores to predict GPA. Data on 78 students in this 7th grade class give us the following statistics regarding these two variables. (IQ is on a scale of 0-140; GPA is on a scale of 0-400.).

IQ	GPA
$\bar{x} = 100$	$\bar{x} = 300$
$s = 10$	$s = 40$
$r = 0.5$	

(a) Find the equation of the least-squares line for this data.

(b) I scored a 80 on my IQ test, what is my predicted GPA?

(a) $b_1 = r \cdot \frac{s_y}{s_x} = (0.5)\frac{40}{10} = 2$

$b_0 = \bar{y} - b_1\bar{x} = 300 - (2)100$

$= 300 - 200 = 100$

$\boxed{\hat{y} = 2x + 100}$ where $y = GPA$

AND $x = IQ$

(b) $y = 2(80) + 100$

$= \boxed{260}$

3. I wanted to test a four-sided die I had made for the purpose of gambling. The die can read 1, 2, 3 or 4. My hope is the die favors the 4. The maker of the die said that a roll has a 50% chance of being a 4 and the other three sides were equally likely.

(a) Write the distribution of the die in table form.

(b) Compute μ and σ from the table.

(a)

x	1	2	3	4
$f(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{2}$

(b)

$$\mu = 1\left(\tfrac{1}{6}\right) + 2\left(\tfrac{1}{6}\right) + 3\left(\tfrac{1}{6}\right) + 4\left(\tfrac{1}{2}\right)$$

$$= 3 \qquad \boxed{\mu = 3}$$

$$\sigma^2 = \sum (x_i - \bar{x})^2 f(x)$$

$$= (1-3)^2 \cdot \tfrac{1}{6} + (2-3)^2 \cdot \tfrac{1}{6} + (3-3)^2 \cdot \tfrac{1}{6} + (4-3)^2 \cdot \tfrac{1}{2}$$

$$= (4)\tfrac{1}{6} + 1 \cdot \tfrac{1}{6} + 0 + \tfrac{1}{2}$$

$$\sigma^2 = \tfrac{4}{3} \qquad \text{so} \quad \boxed{\sigma = \sqrt{\tfrac{4}{3}}}$$

4. I wanted to test a four-sided die I had made for the purpose of gambling. The die can read 1, 2, 3 or 4. My hope is the die favors the 4. The maker of the die said that a roll has a 50% chance of being a 4 and the other three sides were equally likely. So to test the die I rolled the die 3 times and recorded the following numbers: 1, 1, 4

(a) Compute $\bar{x}$.

(b) Compute the variance.

(a) $\bar{x} = \frac{1+1+4}{3} = \frac{6}{3} = \boxed{2}$

(b) $s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$

$= \frac{1}{3-1} \cdot \left[(1-2)^2 + (1-2)^2 + (4-2)^2 \right]$

$= \frac{1}{2} \left[1 + 1 + 4 \right] = 3$

So $\boxed{\text{VARIANCE} = 3}$

AND STANDARD DEVIATION

IS $\sqrt{3}$

5. I want to use my four-sided die with a regular fair four-sided die.

   (a) What are μ and σ from the fair foursided die.

   (b) If I roll the fair four-sided die and the biased four-sided die and sum the numbers on the two dice then what is μ and σ for this sum of the two dice. Note the rolls of the dice are indepenpent.

(a)

y	1	2	3	4	
$f(y)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	
$yf(y)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{3}{4}$	$\frac{4}{4}$	$= \frac{10}{4}$ SUM IS

$$\boxed{\mu_y = \frac{10}{4} = \frac{5}{2}}$$

$$\sigma^2 = \left(1-\frac{5}{2}\right)^2 \cdot \frac{1}{4} + \left(2-\frac{5}{2}\right)^2 \cdot \frac{1}{4} + \left(3-\frac{5}{2}\right)^2 \cdot \frac{1}{4} + \left(4-\frac{5}{2}\right)^2 \cdot \frac{1}{4}$$

$$= \frac{1}{4}\left[2.25 + .25 + .25 + 2.25\right]$$

$$= \frac{1}{4} \cdot 5 = \frac{5}{4} \quad \text{so} \quad \boxed{\sigma_y = \sqrt{\frac{5}{4}}}$$

(b) $Z = X + Y$

FROM PROB (3)

$$\mu_Z = \mu_X + \mu_Y = 3 + \frac{5}{2} = \boxed{5.5}$$

$$\boxed{\mu_Z = 5.5}$$

$$\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2 \quad \text{SINCE } X \perp Y$$

$$\sigma_Z^2 = \frac{4}{3} + \frac{5}{4} = \frac{31}{12} \quad \text{so} \quad \boxed{\sigma_Z = \sqrt{\frac{31}{12}}}$$

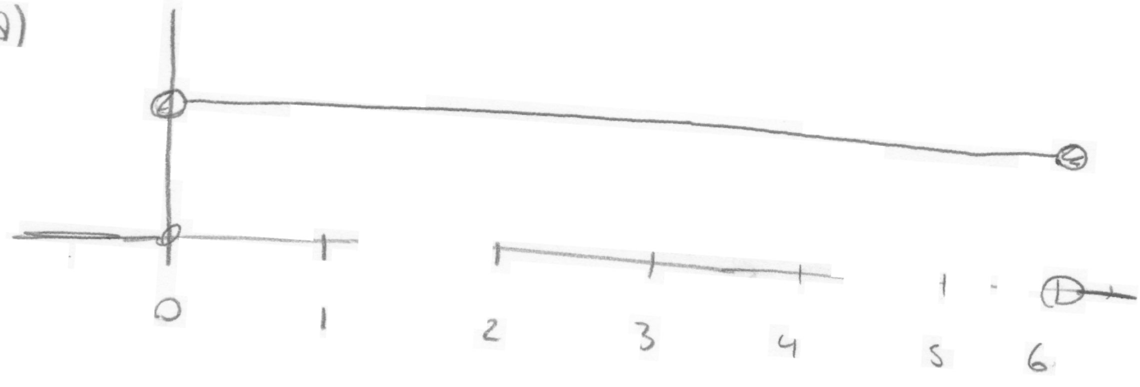
TRY IF X AND Y ARE NOT INDEPENDENT. WHAT INFO. DO YOU NEED?

6. Refer to the sketch below for this question.

(a) What is the height of the given distribution?

(b) Compute probabilty that X is between 1 and 2.

(a)



AREA MUST BE 1 SO HEIGHT = $\boxed{\frac{1}{6}}$

(b) $Pr(1 \le X \le 2) = \frac{1}{6} \cdot (2-1) = \boxed{\frac{1}{6}}$

HEIGHT WIDTH

7. Let X and Y be random variables with means $\mu_X = 2$ and $\mu_{\cancel{Y}} = 2$ and with standard deviations $\sigma_X = 1$ and $\sigma_{\cancel{Y}} = 1$. Assume the $\rho = 0.5$. Define the random variables $A = X - Y$ and $B = 5X$. Compute the mean and variance of A and B.

$A = X - Y$

$$\mu_A = \mu_{X-Y} = \mu_X - \mu_Y = 2 - 2 = 0 \qquad \boxed{\mu_A = 0}$$

$$\sigma_A^2 = \sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2 - 2\rho\,\sigma_X\,\sigma_Y$$

$$= (1)^2 + (1)^2 - 2(.5)(1)(1)$$

$$= 2 - 1 = 1 \qquad \boxed{\sigma_A = 1}$$

$B = 5X$

$$\mu_B = 5\mu_X = 5(2) = 10 \qquad \boxed{\mu_B = 10}$$

$$\sigma_B^2 = \sigma_{5X}^2 = 25(\sigma_X)^2 = 25(1)^2 = 25$$

SO $\boxed{\sigma_B = 5}$

TRY COMPUTING

μ_C AND σ_C where ① $C = 2X - 4Y$

② $C = X^2$

③ $C = (X+Y)^2$

CAN YOU DO ALL THREE OF THESE?

8. Assume I have two RV's representing test scores

- Test 1: X with mean $\mu = 10$ and standard deviation $\sigma = 2$
- Test 2: Y with mean $\mu = 12$ and standard deviation $\sigma = 3$

(a) What is the test score range about the mean that represents 95% of the test takers for each test.

(b) Sarah scored an 11 on Test 1 and Jessica scored a 14 on Test 2. Compute the z-scores and determine who had the higher score (Assume Normal).

(a) $\mu \pm 2 \cdot \sigma$

2 for 95%

Test 1	Test 2
$10 \pm 2(2)$	$12 \pm 2(3)$
6 to 14	6 to 18

(b) SARAH

$$Z = \frac{X - \mu}{\sigma} = \frac{11 - 10}{2} = \frac{1}{2}$$

JESSICA

$$Z = \frac{X - \mu}{\sigma} = \frac{14 - 12}{3} = \frac{2}{3}$$

JESSICA SCORED HIGHER

WHAT WERE THE PERCENTILES THEY SCORED IN?

9. There are many rating systems for students. Often graduate schools and potential employers will ask for a recommendation and then ask the recommender to rate the student on scale; two that I have seen are below. Compare the two scales and discuss any possible bias.

Scale 1	Scale 2
exceptional	top 5%
top 1%	top 10%
top 5%	above average
top 10%	average
top 20%	below average
above average	Not recommmended
Not recommmended	

SCALE 1 leaves too little differentiation between exceptional, top 1% AND top 5%. With approximately 20 students in a class what is the difference between these three?

10. A taxi driver owns two taxis a fuel efficient Prius and a classy luxury car. The driver wants to test whether she gets higher tips in one car verse the other. She decides to drive every day for the the next 60 days. Design an experiment where she can compare the two cars.

The driver should randomize the days

30 days Prius

30 days Lexus

OR

YOU COULD BLOCK WEEKENDS AND WEEKDAYS AND THEN RANDOMIZE THE TAXI CAR SELECTION.